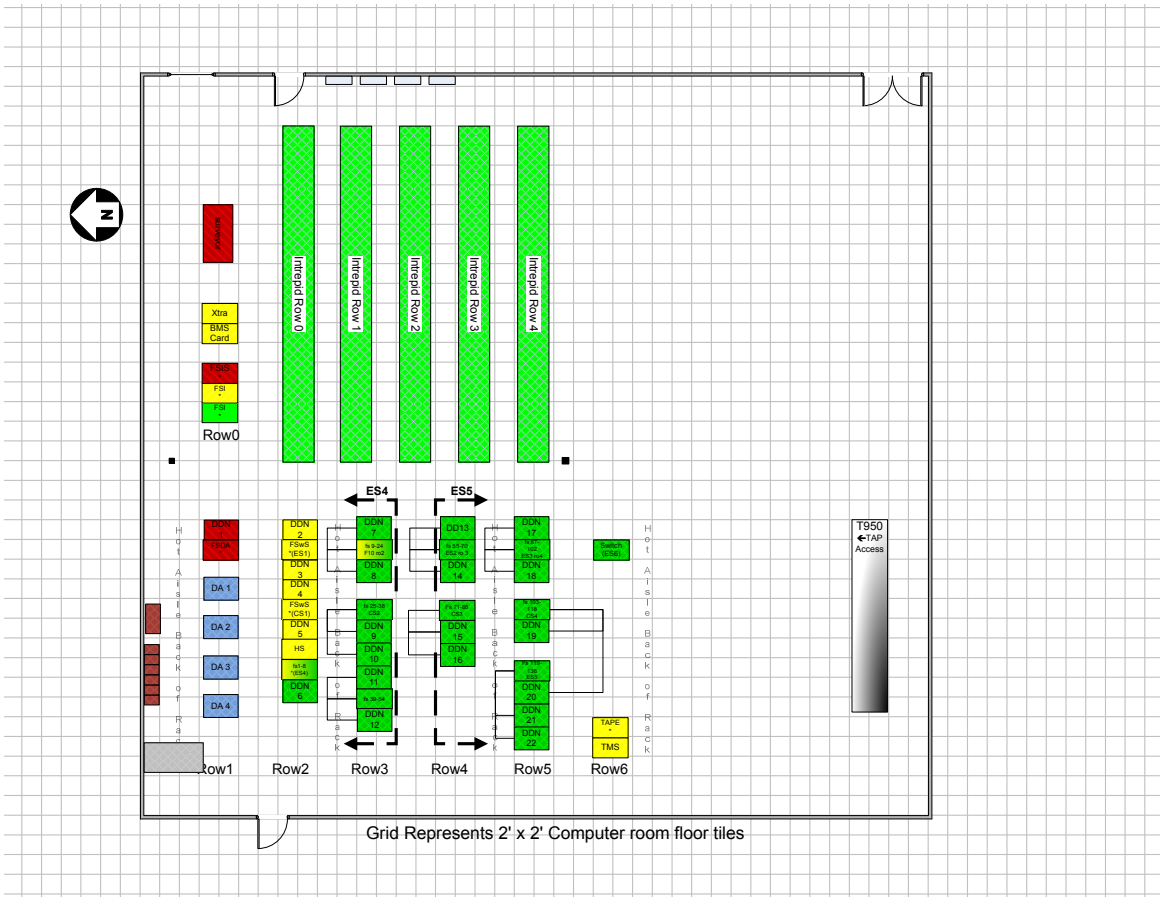
ALCF Infrastructure Getting Started Workshop

William Scullin
Senior High Performance Computing Systems Administrator
Leadership Computing Facility

The World Inside The ISSF



ALCF Resources

INCITE

T&D

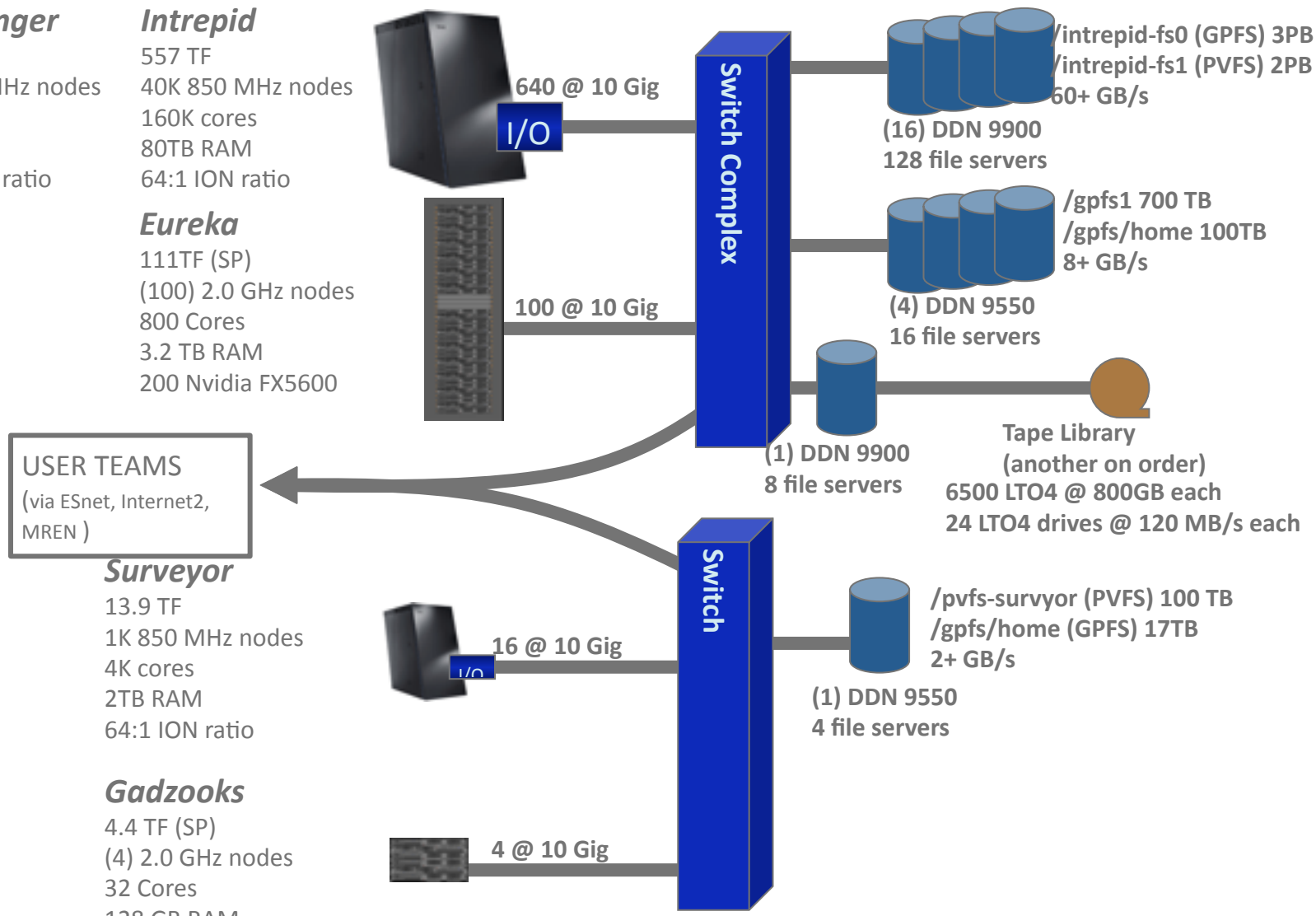
Challenger
13.9TF
1K 850 MHz nodes
4K cores
2TB RAM
16:1 ION ratio

Intrepid
557 TF
40K 850 MHz nodes
160K cores
80TB RAM
64:1 ION ratio

Eureka
111TF (SP)
(100) 2.0 GHz nodes
800 Cores
3.2 TB RAM
200 Nvidia FX5600

Surveyor
13.9 TF
1K 850 MHz nodes
4K cores
2TB RAM
64:1 ION ratio

Gadzooks
4.4 TF (SP)
(4) 2.0 GHz nodes
32 Cores
128 GB RAM
8 Nvidia FX5600



System Details

Blue Gene /P System	Surveyor	Intrepid	Challenger
Function	Test & Development	INCITE Production	INCITE Test & Development
Login address	surveyor.alcf.anl.gov	intrepid.alcf.anl.gov	challenger.alcf.anl.gov
Login OS	SLES 10 SP 3	SLES 10 SP 3	SLES 10 SP 3
Login CPUs	4 PPC970MP @ 2.5 GHz	4 PPC970MP @ 2.5 GHz	4 PPC970MP @ 2.5 GHz
Login memory	4GB	4GB	4GB
BGP CPU (quad core)	850MHz PPC450fp2	850MHz PPC450fp2	850MHz PPC450fp2
# Nodes / # Cores	1024 / 4096	40,960 / 163,840	1024 / 4096
BGP Memory	2TB (2GB per node)	80TB (2GB per node)	2TB (2GB per node)
# I/O nodes (Ratio)	16 @ 10 Gig (64:1)	640 @ 10 Gig (64:1)	64 @ 10 Gig (16:1)
BGP Compute OS	CNK, ZeptoOS, Plan 9	CNK, ZeptoOS	CNK, ZeptoOS

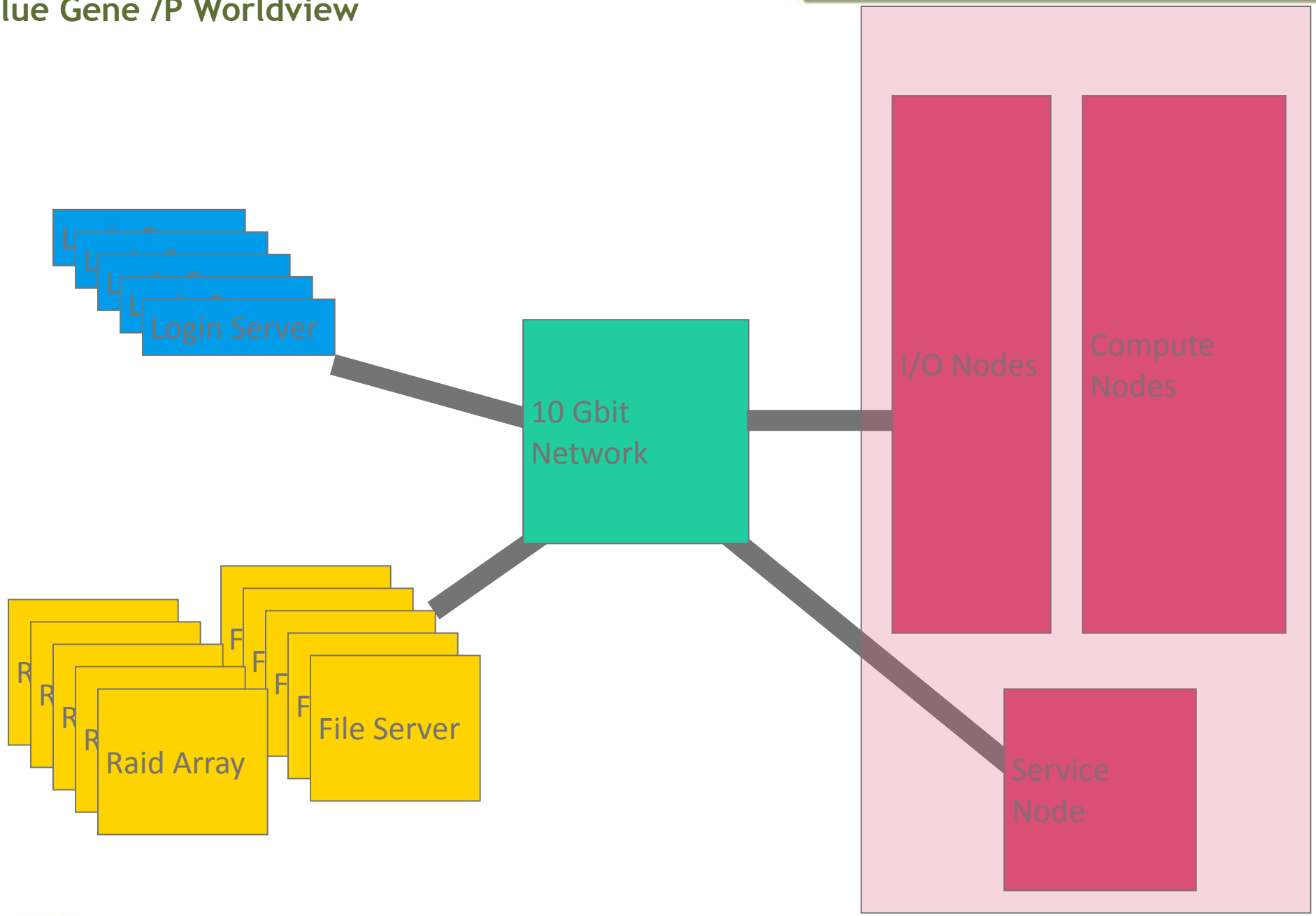
- Login to ALCF Resources is via ssh with cryptocard authentication
- Round robin DNS will place you on a login node named login[1..n].<machine>.alcf.anl.gov
- Logins are for compilation and job submission only:
  - You may use parallel make, but be gentle
  - Do not do I/O to your home directory

System Details

Visualization System	Gadzooks	Eureka
Function	Test & Development	Production INCITE
Login access	gadzooks.alcf.anl.gov	eureka.alcf.anl.gov
Node OS	SLES 10 SP 3	SLES 10 SP 3
Node CPU (quad core)	2 Intel Xeon E5405 @	2 Intel Xeon E5405 @
Nodes / Cores	4 / 32	100 / 800
Memory	128GB (32GB per node)	3.2 TB (32GB per node)
Nvidia FX5600s	8	200
Interconnect	10 GigE (Myrinet)	10 GigE (Myrinet)

- Login to ALCF Resources is via ssh with cryptocard authentication
- Round robin DNS will place you on a login node named login[1..n].<machine>.alcf.anl.gov
- Logins are for compilation and job submission only:
  - You may use parallel make, but be gentle
  - Do not do I/O to your home directory
- Logins are separate from visualization nodes

Blue Gene /P Worldview



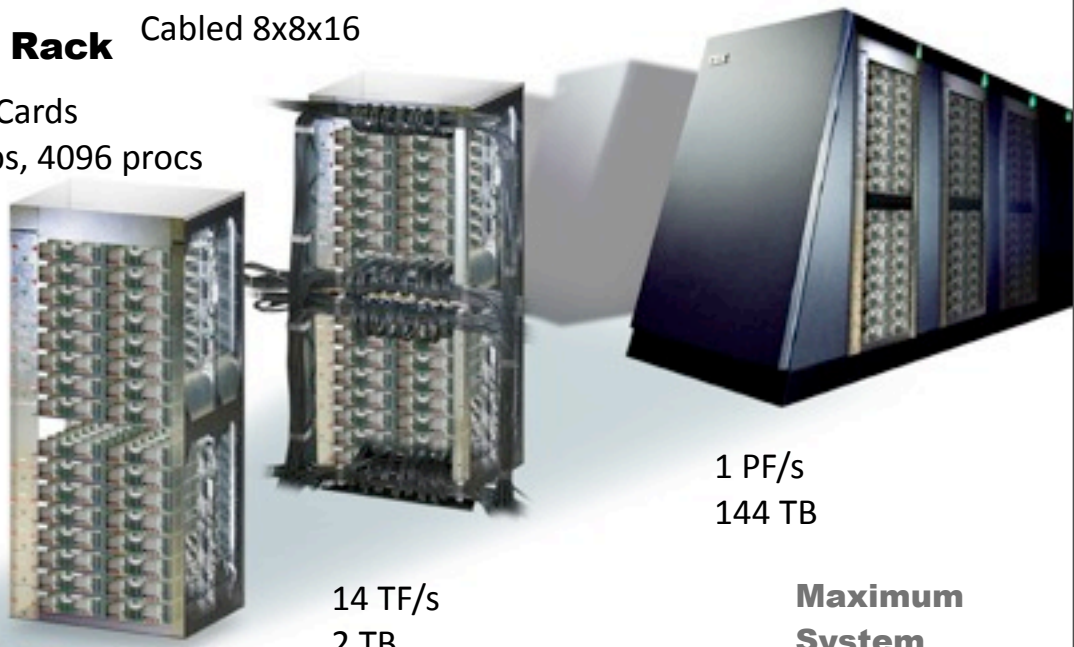
Login node notes

- Intrepid, Challenger, and Surveyor login nodes are showing their age
  - Roughly equivalent to a quad code PowerMac G5
  - People tend to congregate on login1 and login2
    - **Do try another login if things seem slow**
    - **Challenger shares file systems with Intrepid - you can compile there too**
  - The uplink from the logins to the i/o network is a single 10GigE connection
    - Eureka logins have their own 10GigE connection making it a better choice for i/o intensive work, heavy i/o work on the PowerPC logins slows everything down
- Be a good citizen!
  - Don't run parallel make for more cores (4) than a login has
  - Clean up your files when done
    - We purge /tmp, /scratch, and /rscratch, but it's better if you clean up after compiles
  - We will terminate processes impacting stability and other's use of the logins
    - We will contact you if we do
    - We can usually accommodate your workflows – just ask
- Challenger and Intrepid

Blue Gene /P Overview

Intrepid System
40 Racks

Rack
Cabled 8x8x16
32 Node Cards
1024 chips, 4096 procs



1 PF/s
144 TB

Node Card
(32 chips 4x4x2)
32 compute, 0-2 IO cards

14 TF/s
2 TB

Maximum System
256 racks
3.5 PF/s
512 TB

Compute Card
1 chip, 20 DRAMs

435 GF/s
64 GB

Chip
4 processors

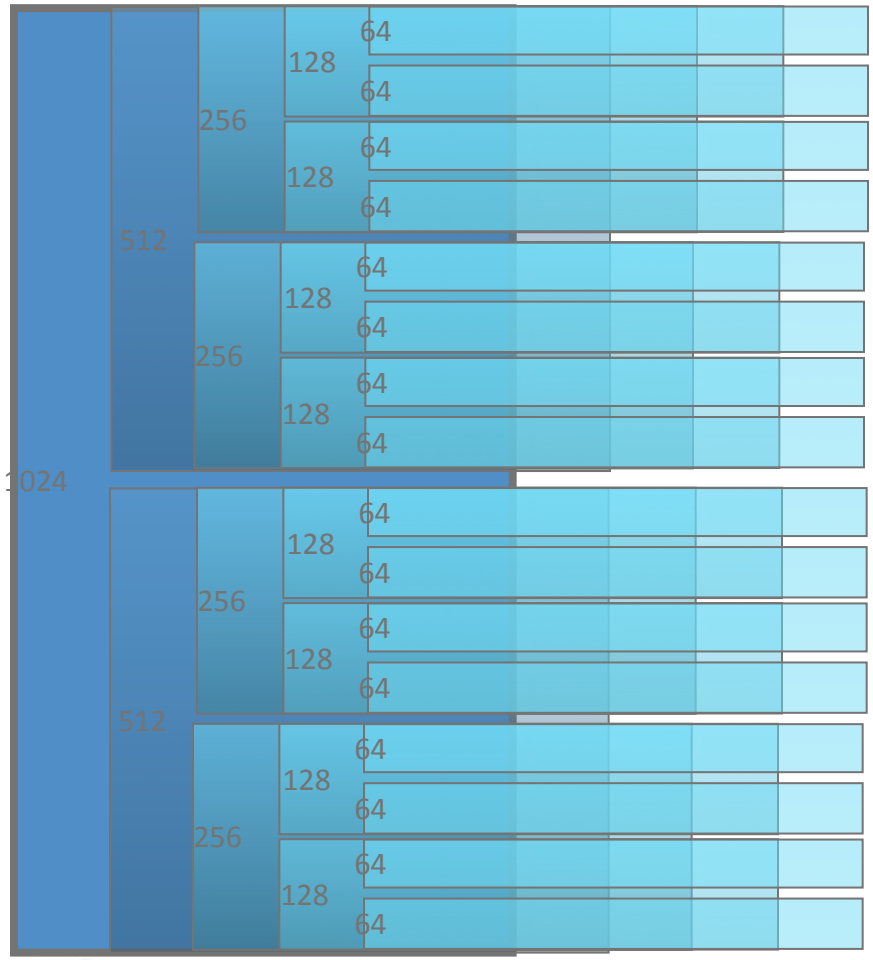
13.6 GF/s
2.0 GB DDR
Supports 4-way SMP

850 MHz
8 MB EDRAM

Front End Node / Service Node
System p Servers
Linux SLES10

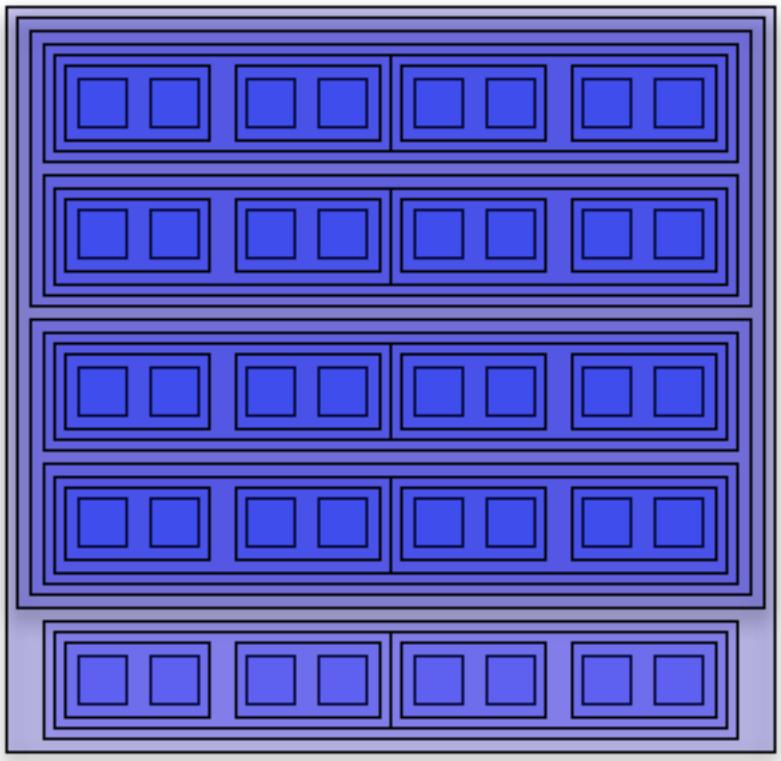
HPC SW:
Compilers
GPFS
ESSL
Cobalt

Blue Gene Single Rack Partitions ("blocks")



- 1 I/O node for each 64 compute nodes, hardwired to specific set of 64
  - *Minimum partition size of 64 nodes*
- Partition sizes: 64, 128, 256, 512, 1024
  - *Any partition < 512 nodes will get a mesh network layout and not a torus.*
  - *Any partition <512 nodes will get a non-optimal I/O tree network.*
  - *Do not do performance testing on <512 nodes*
- Smaller partitions are enclosed inside of larger ones
  - *Not all partitions are available at all times*
  - *Once a job is running on one of the smaller partitions, no jobs can run on the enclosing larger partitions*
- Configuration changes frequently
  - ***partlist** shows partition state*
- Processes are spread out in a pre-defined mapping, alternate and sophisticated mappings are possible

Blue Gene Multiple Rack Partitions ("blocks")



- The following number of large block sizes are possible :
  - 1 40960
  - 1 32768
  - 2 16384
  - 4 8192
  - 9 4096
  - 19 2048
- Not all possible blocks are available at the same time due to wiring dependencies.
- **partlist** will show you if a large free block is busy due to a wiring dependency
- The 40960 node block is generally only available through a reservation, this will change with Challenger online
- One rack, R47, is generally reserved for debugging and testing on Intrepid making only the following blocks possible on R4: 4096, 2 4096, , this will change with Challenger online
- Mesh partitions are available by reservation only

Resource Manger and Job Scheduler

- Cobalt - locally developed open source resource manager and scheduler
  - Uses a “cost function” to compute the priority of a job.
- Used on all systems
- Standard commands (qsub, qstat, qdel, qalter)
- Surveyor queues
  - default: Runs the “unicef” cost function; minimize large job starvation while getting good turnaround times
  - Max runtime is 1 hour, no more than 12 jobs running per user, no more than 20 jobs queued per user.
- Intrepid queues
  - prod: Wfp^3 policy; gives priority to larger jobs; will automatically “drain” the machine.
    - Minimum 512 node jobs, max runtime is 12 hours, no more than 5 jobs running per user, no more than 20 jobs queued per user
  - prod-devel: unicef policy (like surveyor)
    - no minimum job size, max runtime is 1 hour, no more than 5 jobs running per user, no more than 20 jobs queued per user
    - This queue will go away when challenger.alcf.anl.gov becomes publically available

Reservations

- Reservations
  - Should be the exception not the rule see: https://wiki.alcf.anl.gov/index.php/Queuing#Reservations for details
  - Email reservation requests to ***support@alcf.anl.gov***
  - View reservations with **showres**
  - Release reservations with **ureleaseres**
- Special reservations
  - R.pm: Preventative maintenance reservation Mondays from 8am to 8am
    - Typically complete in the early evening
  - R.hw* or R.sw*: Administrative reservation while addressing hardware or software issues
- This workshop will use:
  - **R.gs11**
  - Please pay attention to reservation constraints with regard to this queue

Allocation Management

- Every user must have at least one Project they are assigned to
- Projects are then given allocations
  - Allocations have an amount, start, and end date and are tracked separately; Charges will cross allocations automatically. The allocation with the earliest end date will be charged first, until it runs out, then the next, and so on
- Charges are based on the partition size, NOT the # of nodes or cores used!
- Reservations are charged for the full time they are active
- Will be managed with clusterbase
  - Use the 'cbank' command (see 'cbank --help')
- Examples:

  # list all charges against a particular project
    - cbank -l charge -p <projectname>

  # list all active allocations for a particular project
    - cbank -l allocation -p <projectname>

File systems - Intrepid

- Phase I storage: (4) DDN9550 @ 2.2 GB/s each, (8) fs for home, (16) for GPFS
  - /gpfs/home
    - Intended for source code, binaries, etc.. NOT DATA
    - GPFS, 275TB,
    - Backed up via snapshots and tape
- Local storage on login nodes
  - /scratch is available.
    - XFS, 70GB, relatively fast, temporary
    - NOT mounted on BG/P
    - NOT cross-mounted between nodes
  - /rscratch
    - NFS
    - NOT mounted on BG/P
    - Cross-mounted on all intrepid logins
    - Meant for helping with distcc

File systems - Intrepid

- Phase II storage: (16) DDN9900 @ 5.5 GB/s each, 128 file servers
  - /intrepid-fs0
    - Intended for very fast parallel IO, program input and output
    - GPFS, 4.5 PB, 60+ GB/s
    - Not backed up, but you can initiate archive via HPSS
    - Contains
      - /intrepid-fs0/users/${USER}/scratch
      - /intrepid-fs0/users/${USER}/persistent
      - Moving to project allocations
  - /intrepid-fs1
    - Available only with the kernel profile 'pvfs'
    - Intended for very fast parallel IO, program input and output
    - PVFS, 0.5 PB, 50+ GB/s
    - Not backed up, but you can initiate archive via HPSS
    - Contains
      - /intrepid-fs1/users/${USER}
  - We strongly prefer that users run from /intrepid-fs0 and write to /intrepid-fs0 or /intrepid-fs1. Job I/O to /gpfs/home is viewed as anti-social and is not supported
  - /gpfs/home may be mounted read-only on IONs in the future

File systems - Surveyor

- Phase I storage: (1) DDN9550 @ 2.2 GB/s each, (8) fs for home, (16) for GPFS
  - /gpfs/home
    - Intended for source code, binaries, etc.. NOT DATA
    - GPFS, 15TB,
    - Backed up via snapshots
  - /pvfs-surveyor
    - Intended for fast parallel IO, program input and output
    - PVFS, 88TB
    - Not backed up, CURRENTLY NO TAPE ACCESS
  - We strongly recommend that you avoid writing to /gpfs/home. It is viewed as anti-social and is not supported
- Local storage on login nodes
  - /scratch is available.
    - XFS, 70GB, relatively fast, temporary
    - NOT mounted on BG/P

Backups and Archival

- Backups
  - Snapshots of home directories are done nightly
    ~/.snapshot
  - home directories are also backed up to tape
    - have not had a single restore request from users
  - Data directories will not be backed up
- Archives
  - Archive service is available via HPSS
    - HSI is an interactive client
    - HTAR is great for lots of small files
      - Path name is limited to 155 chars in the prefix and 100 bytes for the name (prefix/name)
      - File size is limited to 64 GB.
  - GridFTP access to HPSS is available
    - Should be significantly faster

Getting data in and out

- GridFTP is also available to move data in and out of the site
  - Other site must accept our CA
  - ssh / cryptocard access available
- Obviously scp is also available
  - If you must use scp, eureka is a better system to scp to from all paths will be the same as they are on Intrepid
  - Eureka is also a better host for compressing and uncompressing large file archives

Mailing Lists

- For each Blue Gene resource there are two mailing lists.
- Visualization resource related announcements are sent to the mailing list of the associated Blue Gene
- <resource>-users@alcf.anl.gov
  - Mandatory, auto-built from all users with active accounts
  - Important announcements impacting the entire community
    - Security issues
    - Major downtimes
    - Policy changes
    - Long-term news
- <resource>-notify@alcf.anl.gov
  - For the active community
  - Operational status announcements
  - Initially subscribed with account creation
  - Subscribe/unsubscribe as you wish

Cyber Security

- Argonne computer user agreements
  - Agreed to at account request time
- Standard Argonne computer security rules apply
  - No sharing accounts
    - We WILL know if, for instance, you are letting your grad student use your account.
  - Acceptable use
  - Etc.
- No passwords are allowed for accessing the systems
  - SSH keys used to access Surveyor
  - CRYPTOCard token required to access Intrepid
    - CRYPTOCard tokens will work for Surveyor as well
- Data policies are available on the web:
  - http://www.alcf.anl.gov/support/usingALCF/docs/dataprivacy.php
  - If you have prohibited data (PII, UNCI, etc.) please contact support@alcf.anl.gov

When things go wrong… logging in

- Check to make sure it's not maintenance
  - Often login nodes on both BG/P and data analytics systems are closed off during maintenance to allow for activities that would impact users
  - There should be a mention in the bi-weekly maintenance announcement and the pre-login banner message
  - An all-clear will be sent out at the close of maintenance
- Check your password
  - Remember that CryptoCARD passwords
    - Require a pin at the start
    - Are all uppercase
    - Are all all hexadecimal characters (0-9, A-F)
  - Try a fresh password
  - Walk through the unlock and resync steps at: https://wiki.alcf.anl.gov/index.php/User_Support#My_Cryptocard_isn.27t_working._How_do_I_proceed.3F
- Connect with **ssh –vvv** and record the screen output, your ip address and hostname, and the time that you attempted to connect. Include this information in any ticket to speed debugging

When things go wrong… running

- Cobalt jobs, by default, produce three files prefixed with either the job number or the name specified with qsub's **–O** option:
  - **$PREFIX.output** with output to standard out from you application
  - **$PREFIX.error** with output from the control system, scheduler, and your application directed to standard error
  - **$PREFIX.cobaltlog** with a record of the environment and the submission command
    - Generated at submit time
    - Very useful for the support team in debugging
- Only cobaltlog is generated at submit time, the others at runtime
- At boot, the .error file will have a non-zero size for non-script jobs
  - Most of the messages are related to booting, it's a decent way to follow startup progress
  - Script jobs leave the handling of standard error up to you, though most do use the .error file
  - We'll walk through a normal boot in a moment…
- If you think there is an issue, it's best to save all three files

When things go wrong… running

- You'll see RAS events appear in your .error file it's not always the sign of trouble
- RAS stands for Reliability, Availability, and Serviceability
- Few are a sign of a serious issue, most are system noise
  - Messages have a severity associated with them
    - INFO
    - WARN
    - ERROR
    - FATAL
  - Only FATAL RAS events will lead to the termination of your application
  - Still worth watching as they may be the sign of an application issue
- The most common RAS event is **APPL_0A2B**
  - Accounts for 58% of all RAS events over the last year
  - an INFO RAS event meaning the DMA unit reception FIFO is full
  - Sign of an application issue that impacts performance
    - Applications are posting too many sends before receives
    - Less severe cases may be addressed through tuning the DCMF_RECFIFO variable size
  - Totally preventable

When things go wrong… running

- Also common, but ignorable
  - **KERN_080A**
    - Correctable single bit errors
  - **KERN_0802**
    - Correctable single symbol error
  - **KERN_1015**
    - correctable torus receiver error
  - **KERN_1021**
    - correctable receiver errors

When things go wrong… running

Remember:

The VelociRAStor says only you can prevent **RAS Events**

Getting Help

Problems or Questions:

Check:

- Getting Started: http://www.alcf.anl.gov/support/usingALCF/usingsystem.php
- ALCF Wiki: https://wiki.alcf.anl.gov/index.php/Main_Page
- ALCF web pages: http://www.alcf.anl.gov
- Intrepid Status: http://status.alcf.anl.gov/intrepid/activity (beta, a.k.a. The Gronkulator)

Contact:

- e-mail: **support@alcf.anl.gov**
- phone: **630-252-3111 (866-508-9181)**
- Your catalyst

Thank you and Questions

Thanks for listening!

Any questions?